Weak Values as Context Dependent Values of Observables and Born's Rule

Akio Hosoya[1, *] and Minoru Koga[1, †]

[1]*Department of Physics, Tokyo Institute of Technology, Tokyo, Japan*
(Dated: January 19, 2013)

We characterize a value of an observable by a 'sum rule' for generally non-commuting observables and a 'product rule' when restricted to a maximal commuting subalgebra of observables together with the requirement that the value is unity for the projection operator of the prepared state and the values are zero for the projection operators of the states which are orthogonal to the prepared state. The crucial requirement is that the expectation value and the variance of an observable should be independent of the way of measurement, i.e., the choice of the maximal commuting subalgebra of observables. We shall call the value a *'contextual value'*. We show that the contextual value of an observable coincides with the weak value advocated by Aharonov and his colleagues by demanding the consistency of quantum mechanics with Kolmogorov's measure theory of probability. This also gives a derivation of Born's rule, which is one of the axioms of conventional quantum mechanics.

I. INTRODUCTION

In the conventional Copenhagen interpretation of quantum mechanics, the value of physical quantity emerges only after measurement but not before. However, the idea of value only after measurement becomes problematic when one considers quantum gravity, because measurements presuppose spacetime which emerges after the quantum era of the Universe apart from the fact that there are no observers who can make projective measurements of the Universe [1]. Independently Ozawa [3] has pursued the possibility of assigning a context dependent value for a physical quantity before measurement, where the context describes outcomes of the measurements to be performed.

On the other hand, Born's rule [4] is an axiom of conventional quantum mechanics, though there is literature which claims that the rule can be derived. It seems that there is a lack of either physical intuition because of a lack of a measurement process or a lack of mathematical rigor. We shall give a brief review of these notions restricting our attention to the orthodox view in the discussion part of the present work but not to the many world interpretation of quantum mechanics [5, 6].

In the present paper we show that the contextual value of an observable coincides with the weak value proposed by Aharonov and his colleagues [7, 8] on the basis of consistency of quantum mechanics with Kolmogorov's measure theory of probability [9] and that this consistency also leads to Born's rule as the unique probability measure of events to be discovered by the post-selection.

The organization of the present paper is as follows. After a brief preliminary in Sect 2, we present in Sect 3 a theorem that the weak values are the contextual values of observables and give a derivation of Born's rule. The final section is a summary and discussion of our results and includes a brief review of the claims of the derivation of Born's rule.

II. PRELIMINARY

We first set up the general framework. Let $\mathcal{N}$ be a set of observables which act on the Hilbert space $\mathcal{H}$ of finite dimension, i.e., $\dim \mathcal{H} = N$. Let $\mathcal{V}(\mathcal{N})_{\max}$ be the set of all maximal abelian subalgebras of $\mathcal{N}$ and choose a maximal abelian subalgebra $V_{\max} \in \mathcal{V}(\mathcal{N})_{\max}$, noting that the choice is not unique. Then we can select a complete orthonormal system $\{|\omega\rangle\}_{\omega \in \Omega}$ as the set of all simultaneous eigenstates of the elements of $V_{\max}$, where $\Omega := \{\omega_1, \cdots, \omega_N\}$. Hereafter, we identify the eigenspace of $|\omega\rangle\langle\omega|$ with ω. Therefore, Ω represents the way of orthogonal decomposition of the Hilbert space $\mathcal{H}$. We call the set $\{|\omega\rangle\}_{\omega \in \Omega}$ 'context' induced by $V_{\max}$ which is chosen from $\mathcal{V}(\mathcal{N})_{\max}$. In due course we will use Aharonov's formalism to interpret the complete orthonormal system $\{|\omega\rangle\}_{\omega \in \Omega}$ as the set of the states to be 'post-selected'. Thus, since we fix the prepared, i.e., pre-selected state to be $|\psi\rangle$ throughout this paper, the 'context' is the context of the measurement to be performed. In the case of a finite dimensional Hilbert space, a different choice

of $V_{\max} \in \mathcal{V}(\mathcal{N})_{\max}$ amounts to a unitary transformation of the complete orthonormal system $\{|\omega\rangle\}_{\omega \in \Omega}$. For infinite dimensional case, we have many types of von Neumann algebras [10, 11].

We introduce the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ adopting the standard notation of Kolmogorov's measure theory of probability [9] as follows. The sample space is $\tilde{\Omega} := \{\omega \in \Omega \mid \langle \omega | \psi \rangle \neq 0\}$, and the σ-field $\tilde{\mathcal{F}}$ over $\tilde{\Omega}$ is the power set of $\tilde{\Omega}$. $\tilde{P}$ is a probability measure over $\tilde{\mathcal{F}}$. The measure $\tilde{P}$ satisfies the relation: $\tilde{P}(\{\omega_1\} \cup \{\omega_2\}) = \tilde{P}(\{\omega_1, \omega_2\}) = \tilde{P}(\{\omega_1\}) + \tilde{P}(\{\omega_2\})$. This means the probability of the event that ω_1 is measured *or* ω_2 is measured is the sum of $\tilde{P}(\{\omega_1\})$ and $\tilde{P}(\{\omega_2\})$. This probability space excludes the case $\{\omega \in \Omega | \langle \omega | \psi \rangle = 0\}$. We shall deal with the above case at the end of Sect 3. For notational simplicity, hereafter we represent the set $\{\omega\} \in \mathcal{F}$ as ω so that $\tilde{P}(\omega) := \tilde{P}(\{\omega\})$. For a fixed choice of $V_{\max}$, the field $\mathcal{F}$ is also fixed provided that we specify the prepared state $|\psi\rangle$.

In what follows, we use the standard notation that $\mathbb{R}$ and $\mathbb{C}$ are the fields of real and complex numbers, respectively, and $\sigma_x, \sigma_y, \sigma_z$ are the Pauli matrices.

III.   THE MAIN THEOREM

We denote the *contextual value* of $A \in \mathcal{N}$ in $\omega \in \tilde{\Omega}$ by $\lambda_\omega(A)$ and define it as the nonzero map $\lambda_\omega : \mathcal{N} \to \mathbb{C}$ which satisfies the following conditions.

(i) Sum Rule:

$$\lambda_\omega(A + B) = \lambda_\omega(A) + \lambda_\omega(B), \quad \forall A, B \in \mathcal{N}$$

(ii) Product Rule:

$$\lambda_\omega(TS) = \lambda_\omega(T)\lambda_\omega(S), \quad \forall T, S \in V_{\max}$$

(iii) Initial condition:

$$\lambda_\omega(|\psi\rangle\langle\psi|) = 1, \qquad \lambda_\omega(|\psi^\perp\rangle\langle\psi^\perp|) = 0,$$

where $|\psi\rangle$ is the prepared (pre-selected) state and $|\psi^\perp\rangle$ is an arbitrary state orthogonal to $|\psi\rangle$.

(iv) Invariance: The expectation value and the variance of an observable A defined by

$$Ex(A) := \sum_{\omega \in \tilde{\Omega}} \tilde{P}(\omega)\lambda_\omega(A), \tag{1}$$

$$Var(A) := \sum_{\omega \in \tilde{\Omega}} \tilde{P}(\omega)|\lambda_\omega(A) - Ex(A)|^2 \tag{2}$$

are independent of the choice of $V_{\max}$ and depends only on the prepared state $|\psi\rangle$ and A.

Some remarks on these conditions are in order. The sum rule (i) excludes the possibility that the values $\lambda_\omega(A)$ and $\lambda_\omega(B)$ are eigenvalues of the observables A and B, if they do not commute. A simple example of this is σ_x and σ_y with the eigenvalues ± 1, while the sum $\sigma_x + \sigma_y$ has eigenvalues $\pm\sqrt{2}$. Historically, the rule (i) was introduced by von Neumann [10] when he made an argument to refute the hidden variable theory. The product rule (ii) means that the map λ_ω is a character on $V_{\max}$ and implies that it can be seen as a '*valuation*' on $V_{\max}$ [2]. This means that the contextual value for an observable $T \in V_{\max}$ can be identified with a classical value of T. We are going to generalize it to a contextual value of an observable $A \notin V_{\max}$, which has no counterparts in classical theory. We see that $\lambda_\omega(\mathbf{1}) = 1$ from (ii). The requirement (iii) is just an assumption based on an intuition that the value of the projection operator of the initial state $|\psi\rangle$ is unity and the values of the projection operators of the states $|\psi^\perp\rangle$ which are orthogonal to the prepared $|\psi\rangle$ are zero if we know definitely what the initial states is. The invariance of the expectation value and the variance of an obsevable A in the requirement (iv) uniquely determines the normal distribution function of the observable obtained after repeating many experiments according to the central limit theorem.

From the sum rule (i) we can apply Riez's theorem and using the normalization $\lambda_\omega(\mathbf{1}) = 1$, we can write the contextual value of $A \in \mathcal{N}$ as

$$\lambda_\omega(A) = \frac{\text{Tr}[W_\omega A]}{\text{Tr}[W_\omega]}, \tag{3}$$

with W_ω being some operator. The product rule (ii) for the maximal abelian subalgebra $V_{\max}$ narrows down the expression for the operator W_ω to

$$W_\omega = a|\alpha\rangle\langle\omega| + b|\omega\rangle\langle\beta| + W_\omega^\perp, \quad a, b \in \mathbb{C}, \tag{4}$$

for some $|\alpha\rangle \in \mathcal{H}$ and $\langle\beta| \in \mathcal{H}^*$, where $\mathcal{H}^*$ is the dual space of $\mathcal{H}$. Here $W_\omega^\perp$ is an operator which satisfies $\{W_\omega^\perp, |\omega\rangle\langle\omega|\} = 0$. The requirement (iii) implies that W_ω has a form $W_\omega = |\psi\rangle\langle q| + |r\rangle\langle\psi|$ for some $\langle q|$ and $|r\rangle$. Equating this with (4), we find that $|\alpha\rangle = |\psi\rangle$, $\langle\beta| = \langle\psi|$ and $W_\omega^\perp = 0$ so that

$$W_\omega = a|\psi\rangle\langle\omega| + b|\omega\rangle\langle\psi|, \quad a, b \in \mathbb{C}, \tag{5}$$

and the contextual value becomes

$$\lambda_\omega(A) = \frac{a\langle\omega|A|\psi\rangle + b\langle\psi|A|\omega\rangle}{a\langle\omega|\psi\rangle + b\langle\psi|\omega\rangle}. \tag{6}$$

So far our discussion for the value of a physical quantity has not touched upon probabilistic interpretation. At this stage we formally introduce the concept of probability following the standard probability theory of Kolmogorov [9]. We use the formula for the expectation value of an observable $A \in \mathcal{N}$,

$$Ex(A) = \sum_{\omega\in\tilde{\Omega}} \tilde{P}(\omega)\lambda_\omega(A), \tag{7}$$

where the contextual value $\lambda_\omega(A)$ of A in $\omega \in \tilde{\Omega}$ is given by (6). The expectation value $Ex(A)$ is an average of the contextual value $\lambda_\omega(A)$ over the context of measurement $\{|\omega\rangle\}_{\omega\in\Omega}$. The $\lambda_\omega(A)$ corresponds to the random variable for an observable A in the probability theory. Each $\lambda_\omega(A)$ depends on $|\omega\rangle$ and therefore on how we choose the context $V_{\max}$, but the averaging washes out the dependence on the choice of the context, i.e., how we choose the set of the states to be post-selected. We shall show that the condition (iv) leads to a further specification of W_ω, i.e., $b = 0$ in (5) so that

$$W_\omega = |\psi\rangle\langle\omega|. \tag{8}$$

We now state and prove our main result.

Theorem. *A map $\lambda_\omega : \mathcal{N} \to \mathbb{C}$ satisfying the above conditions (i)-(iv) is of the form*

$$\lambda_\omega(A) = \frac{\langle\omega|A|\psi\rangle}{\langle\omega|\psi\rangle}, \text{ for } \omega \in \tilde{\Omega}, \tag{9}$$

and the probability measure is

$$\tilde{P}(\omega) = |\langle\omega|\psi\rangle|^2. \tag{10}$$

The contextual value (9) is identical to the weak value of Aharonov with $|\psi\rangle$ and $\langle\omega|$ being the pre-selected state and the post-selected state, respectively. By this identification, we explicitly see that the context can be interpreted as a set of post-selected states. Note that the probability measure P is a matter of choice in the classical probability theory, while Born's rule (10) is mandatory in quantum mechanics and even its non-negativity is an important aspect of the consequence. The asymmetry of $\langle\omega|$ and $|\psi\rangle$ in (8) comes from the expression for the expectation value (7) in which the post-selected states are summed up on the basis of an intuition that we can choose a pre-selected state but its outcome is uncertain. The arrow of time is built-in in (7). The expression (7) for the expectation value with (8) and (10) reproduces the standard quantum mechanical expectation value,

$$Ex(A) = \langle\psi|A|\psi\rangle, \tag{11}$$

which certainly does not depend on the context $\{|\omega\rangle\}_{\omega\in\Omega}$. The apparent context dependence of $Ex(A)$ in (7) goes away by the completeness relation on $|\psi\rangle$: $\sum_{\omega\in\tilde{\Omega}}\langle\psi|\omega\rangle\langle\omega|A|\psi\rangle = \langle\psi|A|\psi\rangle$. It is instructive to think about the possibility $\tilde{P}(\omega) = |\langle\omega|\psi\rangle|^4$ instead of (10). Then we would have $Ex(A) = \sum_{\omega\in\tilde{\Omega}}|\langle\omega|\psi\rangle|^4 \frac{\langle\omega|A|\psi\rangle}{\langle\omega|\psi\rangle} = \sum_{\omega\in\tilde{\Omega}}|\langle\omega|\psi\rangle|^2\langle\psi|\omega\rangle\langle\omega|A|\psi\rangle$. The context dependence would not go away. We can follow the exact parallel way also for the variance $Var(A)$ to confirm its context independence.

Proof. Consider a functional

$$\mathcal{L}_E(\langle\omega|,\mu) := Ex(A) - \sum_{i=0}^{N-1} \mu_i\big[\sum_{\omega\in\tilde{\Omega}}\langle\psi_i|\omega\rangle\langle\omega|A|\psi\rangle - \langle\psi_i|A|\psi\rangle\big], \qquad \mu_i\in\mathbb{R}, \qquad (12)$$

where μ_i are the Lagrange multipliers and $\{|\psi_i\rangle\}$ is a complete orthonormal system which contains $|\psi\rangle = |\psi_0\rangle$ as a unit vector. The Lagrange multiplier terms give constraints

$$\sum_{\omega\in\tilde{\Omega}}\langle\psi_i|\omega\rangle\langle\omega|A|\psi\rangle - \langle\psi_i|A|\psi\rangle = 0, \ \ \forall\, i, \ \ \forall A\in\mathcal{N}. \qquad (13)$$

Since $\{|\psi_i\rangle\}$ is a complete orthonormal system and $A\in\mathcal{N}$ is arbitrary, the above constraints are equivalent to the completeness relation $\sum_{\omega\in\Omega}|\omega\rangle\langle\omega| = 1$ for $\{|\omega\rangle\}_{\omega\in\Omega}$ as we see comparing with its matrix elements with (13) .

We demand that the variation of $\mathcal{L}_E(\langle\omega|,\mu)$ with respect to $\langle\omega|$ should vanish. That is,

$$\frac{\delta\mathcal{L}_E(\langle\omega|,\mu)}{\delta\langle\omega|} = \frac{\partial\tilde{P}(\omega)}{\partial\langle\omega|}\frac{a\langle\omega|A|\psi\rangle + b\langle\psi|A|\omega\rangle}{a\langle\omega|\psi\rangle + b\langle\psi|\omega\rangle} + \tilde{P}(\omega)\frac{aA|\psi\rangle}{a\langle\omega|\psi\rangle + b\langle\psi|\omega\rangle}$$

$$- \ \tilde{P}(\omega)\frac{a\langle\omega|A|\psi\rangle + b\langle\psi|A|\omega\rangle}{(a\langle\omega|\psi\rangle + b\langle\psi|\omega\rangle)^2}a|\psi\rangle - \sum_i\mu_i\langle\psi_i|\omega\rangle A|\psi\rangle = 0. \ \ \forall A\in\mathcal{N}$$

A particular choice $A = |\psi\rangle\langle\psi|$ implies that $\frac{\partial\tilde{P}(\omega)}{\partial\langle\omega|} \propto |\psi\rangle$, since the second, third and fourth terms are then $\propto |\psi\rangle$. To see the implication of the component of the above equation perpendicular to $|\psi\rangle$ consider the case $A = |\psi^\perp\rangle\langle\psi| + |\psi\rangle\langle\psi^\perp|$, where $|\psi^\perp\rangle$ is a state orthogonal to $|\psi\rangle$ to find that the second and the forth terms cancel each other. We therefore see that the sum of the first and the third terms and the sum of the second and the fourth terms separately vanish, i.e.,

$$\frac{\partial\tilde{P}(\omega)}{\partial\langle\omega|} - \frac{\tilde{P}(\omega)}{a\langle\omega|\psi\rangle + b\langle\psi|\omega\rangle}a|\psi\rangle = 0, \qquad (14)$$

$$\tilde{P}(\omega)\frac{a}{a\langle\omega|\psi\rangle + b\langle\psi|\omega\rangle} - \sum_i\mu_i\langle\psi_i|\omega\rangle = 0. \qquad (15)$$

Adding these two equations yields

$$\frac{\partial\tilde{P}(\omega)}{\partial\langle\omega|} = \sum_i\mu_i\langle\psi_i|\omega\rangle|\psi\rangle \qquad (16)$$

and its integration gives $\tilde{P}(\omega) = \sum_i\mu_i\langle\psi_i|\omega\rangle\langle\omega|\psi\rangle + \tilde{P}_0$, where $\tilde{P}_0$ is an integration constant. Therefore, (15) becomes

$$\big(\sum_i\mu_i\langle\psi_i|\omega\rangle\langle\omega|\psi\rangle + \tilde{P}_0\big)\frac{a}{a\langle\omega|\psi\rangle + b\langle\psi|\omega\rangle} - \sum_i\mu_i\langle\psi_i|\omega\rangle = 0, \forall\omega\in\tilde{\Omega}. \qquad (17)$$

This implies that $\tilde{P}_0 = 0$ and $b = 0$. We note that $\frac{\delta\mathcal{L}_E(\langle\omega|,\mu)}{\delta|\omega\rangle} = 0$ is automatically satisfied. Therefore, the context invariance of the expectation value $Ex(A)$ leads to expressions for the probability measure and the contextual value of an observable A as

$$\tilde{P}(\omega) = \sum_i\mu_i\langle\psi_i|\omega\rangle\langle\omega|\psi\rangle, \qquad (18)$$

$$\lambda_\omega(A) = \frac{\langle\omega|A|\psi\rangle}{\langle\omega|\psi\rangle}. \qquad (19)$$

We further demand that the variance $Var(A) = \sum_{\omega\in\tilde{\Omega}}\tilde{P}(\omega)|\lambda_\omega(A)|^2$ is context independent. To simplify the discussion we set $Ex(A) = 0$ or replace A by $A - Ex(A)$. It is possible to set up a similar variational problem for $Var(A)$ to the previous one for $Ex(A)$. Instead we simply insert the result of (19) into the expression for $Var(A)$ to obtain

$$
\begin{aligned}
Var(A) \quad &= \sum_{\omega\in\tilde{\Omega}}\tilde{P}(\omega)|\lambda_\omega(A)|^2 \\
&= \sum_{\omega\in\tilde{\Omega}}\sum_i\mu_i\langle\psi_i|\omega\rangle\langle\omega|\psi\rangle|\frac{\langle\omega|A|\psi\rangle}{\langle\omega|\psi\rangle}|^2 \\
&= \sum_{\omega\in\tilde{\Omega}}\sum_i\mu_i\frac{\langle\psi_i|\omega\rangle}{\langle\psi|\omega\rangle}\langle\omega|A|\psi\rangle\langle\psi|A|\omega\rangle.
\end{aligned} \qquad (20)
$$

This is independent of the context $\{|\omega\rangle\}_{\omega\in\Omega}$ if and only if $\mu_i = 0$ for $i \neq 0$. The remaining parameter μ_0 is fixed to be unity by the normalization condition $\sum_{\omega\in\tilde{\Omega}}\tilde{P}(\omega) = \mu_0 = 1$. Then we arrive at the main theorem (9) and (10). ∎

Now we can naturally extend the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ to $(\Omega,\mathcal{F},P)$ as follows. The sample space is Ω, and the σ-field $\mathcal{F}$ over Ω is the power set of Ω. P is the probability measure over $\mathcal{F}$ satisfying the following conditions: $P|_{\tilde{\mathcal{F}}} = \tilde{P}, P|_{\mathcal{F}\backslash\tilde{\mathcal{F}}} = 0$. From the theorem, we can write $P(\omega) = |\langle\omega|\psi\rangle|^2$. Thus the probability space $(\Omega,\mathcal{F},P)$ describes the standard quantum mechanical statistics in the context $\{|\omega\rangle\}_{\omega\in\Omega}$.

IV. SUMMARY AND DISCUSSIONS

We have shown that the contextual value of an observable is the weak value by demanding the consistency of quantum mechanics with Kolmogorov's measure theory of probability in conjunction with the consideration of the values of obsevables. The crucial requirement was that the expectation value and the variance of an observable should be independent of the way of measurement. This leads eventually to Born's rule in quantum mechanics. The assumptions of the sum rule (i) and the product rule (ii) may not have much problem and the invariance (iv) of the expectation value and the variance is crucial to relate the contextual value to the concept of probability. The third assumption (iii) is debatable. At the moment we do not have complete justification. We believe the usefulness of (iii) to characterize the contextual value.

Going back to the original motivation in the introduction, we would like to examine in what sense the weak value can be regarded as the value before projective measurements. Let $A(t)$ be an operator at time $t \leq T$ in the Heisenberg representation, where T is the time of the projective measurement. We choose $A(T)$ as an element of $V_{\max}$ so that the post-selected state $\langle a|$ is an eigenstate of $A(T)$, $\langle a|A(T) = \langle a|a$ of the eigenvalue a. The weak value

$$\frac{\langle a|A(t)|\psi\rangle}{\langle a|\psi\rangle} \tag{21}$$

can be interpreted as the value before projective measurement provided that $\langle a|$ is not orthogonal to the pre-selected state $|\psi\rangle$. At $t = T$, the weak value (21) coincides with the eigenvalue a. As stressed before the weak value depends on the context, i.e., the states to be post-selected. The price is that the contextual value is complex in general.

The extension of the derivation of Born's rule to the mixed state case is straightforward but technically involved so that we postpone it for future publication.

From the main theorem we see that the value of an observable is given by the weak value (9). As is by now well known, the weak value is experimentally accessible [12–21] and theoretically analyzed by many people [22–25]. The weak value coincides with one of eigenvalues when the observable is restricted to the maximal abelian subalgebra $V_{\max} \in \mathcal{V}(\mathcal{N})_{\max}$. For a general $A \in \mathcal{N}$, the weak value can be deduced from the information associated with the state to be post-selected, which is circumstantially inferred from the set of eigenvalues to be obtained by projective measurements of all the elements of $V_{\max}$.

The Kochen-Specker theorem [26] tells us that it is not possible to assign non-contextual values to all observables in quantum mechanics. Since the only non-contextual values are eigenvalues of the observables, this theorem does not apply to the weak values. Actually the weak value is explicitly dependent on the context defined by the post-selected states.

In the proof of Gleason's theorem [27], he demanded that the measure over the set of all the subspaces of the Hilbert space $\mathcal{H}$ should be independent of the choice of basis in order to conclude that the measure has a quadratic form in the state vector components if $\dim(\mathcal{H}) \geq 3$. This leads to the standard expression for the expectation value. There is a gap between the probability and the measure over the set of all the subspaces of the Hilbert space $\mathcal{H}$, so that the derived Born's rule lacks the probabilistic interpretation. In a mathematical sense, Gleason's theorem is close to ours. We believe that we have filled the gap by introducing the physical concept of the contextual value of observables defined by the prepared state and the states to be post-selected. A rather surprising finding in the present work is that the contextual value coincides with the weak value.

Zurek [28] claimed that he derived Born's rule from "the environment assisted invariance". He considered an entangled state

$$|\psi\rangle = \frac{|s_1\rangle|e_1\rangle + |s_2\rangle|e_2\rangle}{\sqrt{2}}, \tag{22}$$

where $|s_1\rangle$ and $|s_2\rangle$ are the orthonormal basis of the system, while $|e_1\rangle$ and $|e_2\rangle$ are the orthonormal basis of the environment. Consider an observable: $A \otimes \mathbf{1}$ with $A = |s_1\rangle\langle s_1| - |s_2\rangle\langle s_2|$. The elementary events are $x_1 := \{\langle s_1|\langle e_1|\}$ and $x_2 := \{\langle s_2|\langle e_2|\}$. The sample space is $\Omega' := \{x_1, x_2\}$, and the σ-field is the power set of Ω'. We would like to show that the probabilities to obtain x_1 and x_2 are $1/2$. To see this we compute the weak values: $\lambda_{x_1}(A) = 1$, $\lambda_{x_2}(A) = -1$. We apply the formula (7) noting that the left hand side is zero because of symmetry under the SWAP operation of the states of 1 and 2 both for the system and environment. The operator $A \otimes \mathbf{1}$ does not distinguish the state of environment so that the expectation value has to be symmetric under the SWAP of the system state. Then we see that $P(x_1) - P(x_2) = 0$ on the right had side and arrive at the equal probability $P(x_1) = P(x_2) = 1/2$, using $Ex(\mathbf{1}) = P(x_1) + P(x_2) = 1$. In this simple case, only the discrete symmetry is sufficient to derive Born's rule. Here we see that the symmetry principle should apply to the expectation value to obtain Born's rule. The role of entanglement with the environment is not clear for us, though we suspect a possible connection to the post-selection.

Acknowledgments

[1] C. Isham and A. Doering, *J. Math. Phys.* **49**, 053518 (2008).
[2] A. Döring, arXiv:quant-ph/1004.3573 (2010).
[3] M. Ozawa, Quantum Reality and Measurement: A Quantum Logical Approach, to appear in Foundations of Physics, DOI:10.1007/s10701-010-9462-y and references therein.
[4] M. Born, *Zeitschrift für Physik*, **37** 863 footnote of p865 (1926),(German)
[5] B. DeWitt and R. Graham, eds,*The Many-Worlds Interpretation of Quantum Mechanics* (Princeton Series in Physics, Princeton University Press 1973).
[6] D. Deutsch, *Proc.Roy.Soc.London A* **455**, 3129 (1999).
[7] Y. Aharonov, D. Z. Albert, and L. Vaidman, *Phys. Rev. Lett.* **60**, 1351 (1988).
[8] Y. Aharonov and D. Rohrlich, *Quantum Paradoxes* (Wiley-VCH, Weibheim, 2005).
[9] A. N. Kolmogorov, "Foundation of the Theory of Probability", second English translation edited by N. Morrison with an added bibliography, (A.D. Barcha-Reid University).
[10] J. von Neumann, *Mathematische Grundlagen der Quantumechanik* (Springer, Berlin, 1932), [*Mathematical foundations of quantum mechanics* Ch. III (Princeton University Press, Princeton, 1955).] For criticism, see J.S. Bell, *Rev.Mod.Phys.* **38**, 447 (1966).
[11] M. Takesaki, *Theory of Operator Algebras I, Ch.5*, (Springer,2002)
[12] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, *Phys. Rev. Lett.* **66**, 1107 (1991).
[13] A. Parks, D. Cullin, and D. Stoudt, *Proc. R. Soc. A* **454** 2997 (1998).
[14] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, *Phys. Lett. A* **324**, 125 (2004).
[15] G. L. Pryde *et al.*, *Phys. Rev. Lett.* **94**, 220405 (2005).
[16] Q. Wang, F. Sun, Y. Zhang, J. Li, Y. Huang and G. Guo, *Phys. Rev. A* **73** 023814 (2006).
[17] K. J. Resch, *Science* **319**, 733 (2008).
[18] O. Hosten and P. Kwiat, *Science* **319**, 787 (2008).
[19] K. Yokota, T. Yamamoto, M. Koashi and N. Imoto, *New J. Phys.* **11** 033011 (2009).
[20] A. Palacios-Laloy, F. Mallet, F. Nguyen, P. Bertet, D. Vion, D. Esteve and A. N. Korotkov, *Nat. Phys.* **6** 442 (2010).
[21] N. Brunner and C. Simon, *Phys. Rev. Lett.* **105** 010405 (2010).
[22] G. Mitchison, R. Jozsa, and S. Popescu, *Phys. Rev. A* **76**, 060302(R) (2007).
[23] R. Jozsa, Phys. Rev. A **76**, 044103 (2007).
[24] Y. Shikano and A. Hosoya, *Phys. A:Math. Theor.* **43**, 025304 (2010).
[25] A. Hosoya and Y. Shikano, *Phys. A:Math. Theor.* **43**, 385307 (2010).
[26] S. Kochen, E. Specker, *Journal of Mathematics and Mechanics* **17**59 (1967).
[27] A. M. Gleason, *Journal of Mathematics and Mechanics* **6**, 885 (1957).
[28] W. H. Zurek, *Phys. Rev. Lett.* **90**, 120404 (2003).